Applying Process Mining in Blockchain Transactions

Massimiliano Sampaolo

January 22, 2019

Supervisor: Barbara Re
Second supervisor: Andrea Morichetta

Objective and Motivations

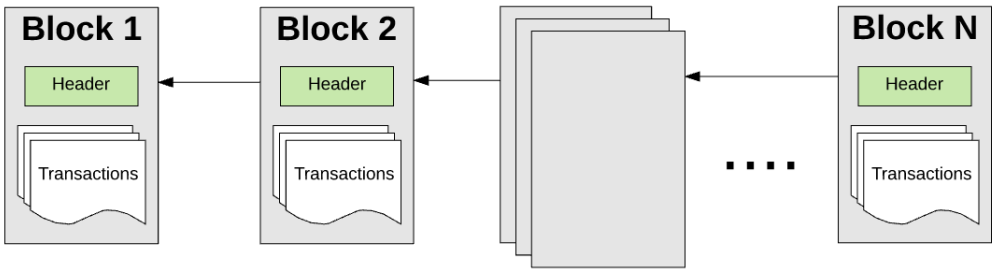
The thesis want to understand and define bindings between **blockchain** transactions using **process mining**.

Why?

- Blockchain and process mining are hot topics in industry and academia
- Infer and analyse the behaviour of systems to increase their quality
- Give a graphical representation of processes defined by software systems

Blockchain

Chain of blocks:



Blockchain characteristics:

- Decentralization
- Trasparency
- Security
- Immutability

Ethereum: EVM, Smart Contracts and DAPP

Business Process Management

A **Business Process** is a collection of related and structured activities undertaken by one or more organisations in order to pursue some particular goal.

Business Process Management (BPM) is the set of activities needed to define, optimize and monitor business processes in order to make effective company's business.

BPMN (Business Process Model and Notation) is a standard language to grapichally represent process models.

Process Mining

The idea is to discover, monitor and improve real processes by extracting knowledge from event logs readily available in today's information systems.

Three form of process mining:

- Process discovery
- Conformance checking
- Enhancement

Used discovery algorithms: **Heuristic Miner**, **Inductive Miner**, **Split Miner**

Heuristic Miner

It mines the control-flow perspective of a process model. It extends alpha algorithm by considering the frequency of traces in the log.

It consists in a set of steps:

- the identification of all the activities
- find the connection between activities and their frequencies
- dependency relation matrix is calculated
- final net is built from the matrix previously created

Inductive Miner

It uses a divide-et-conquer approach:

- builds a directly follow graph (DFG)
- filters infrequent directly-follows dependencies
- find the dominant operator and apply relative cut
- a process tree is generated from each portion

The third step is applied recursively until no more cuts are found.

Split Miner

Split Miner produces a BPMN model in five steps:

Process mining tools

Different software tools are available to support process mining techniques.

The more used are:

- ProM
- Apromore
- Disco

Case Studies

The methodology used for the case studies consists of several steps:

- Retrieving of transaction list from Ethereum
- Generation of log file starting from the transaction list
- Analysis of the log with discovery algoritms previously introduced
- Quality check of the results obtained

RoToHive

RotoHive is a new type of fantasy sports site that runs weekly tournaments.

Results obtained:

Algorithm	**Fitness**	**Precision**	**Generalization**
Split Miner	1	0.20453	0.99892
Inductive Miner	0.99995	0.60294	0.99496
Heuristic Miner	0.99940	0.49889	0.99889

Lessons learned

Sound models discovered and pretty good quality parameters measured.

The three algorithms obtained similiar results. How well an algorithm fits a specific application domain depends from the domain itself regardless the fact that it uses the Blockchain.

The analysis infer the logic of the system. Can be used to increase solution quality, or understand how users interact with a product.

Design and implementation

The system designed recreates the methodology used in the case studies analysis.

The architecture of the Mining Framework: